95% prediction regions: multivariate uncertainty quantification for retrieved atmospheric states

Noel A. Cressie
University of Wollongong, ncressie@uow.edu.au

Sandy Burden
University of Wollongong, sburden@uow.edu.au

95% prediction regions: multivariate uncertainty quantification for retrieved atmospheric states

Abstract

Poster presentation

Keywords

states, regions, multivariate, 95, uncertainty, prediction, quantification, retrieved, atmospheric

Disciplines

Engineering | Science and Technology Studies

Publication Details

N. Cressie & S. Burden (2015). 95% prediction regions: multivariate uncertainty quantification for retrieved atmospheric states. Poster presented at Eleventh International Workshop on Greenhouse Gas Measurements from Space, Pasadena, CA, USA, 16-18 Jun 2015.

95% Prediction Regions: Multivariate Uncertainty Quantification for Retrieved Atmospheric States*

NIASRA
NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

UNIVERSITY OF WOLLONGONG AUSTRALIA

Noel Cressie◇ **and Sandy Burden**

University of Wollongong

ncressie@uow.edu.au sburden@uow.edu.au

◇also Distinguished Visiting Scientist, JPL

*This research was partially supported by NASA grant NNH11ZDA001N-OCO2 (OCO-2 Science Team).

1. Introduction

Retrieving the state of the atmosphere x, including carbon dioxide at different pressure levels, from remote sensing radiance measurements y, is a challenging statistical problem. Simulation experiments are one of the tools used to test, validate, compare, and improve estimates (i.e., predictions) of the multivariate atmospheric state. It is vital to have effective figures of merit (FOMs) to summarise the multivariate prediction-error distribution from these experiments, as these are used to assess the statistical properties of the predicted state variables and to infer prediction intervals for the true values of the state variables.

Consider a simulation experiment where a state-space model is used to generate L independent replications of the state, $x^{(1)},...,x^{(L)}$; data, $y^{(1)},...,y^{(L)}$; predictor, $\hat{x}(y^{(1)}),...,\hat{x}(y^{(L)})$; and prediction error, $\hat{x}(y^{(1)})-x^{(1)},...,\hat{x}(y^{(L)})-x^{(L)}$. The predictor $\hat{x}(y)$ is typically highly non-linear, so the statistical properties of the prediction error are difficult to obtain analytically but can be obtained from the simulation experiment. That is, from the simulations we estimate FOMs and make inference on x using the prediction-error distribution,

$$\hat{x}(y)-x \sim \text{Dist}(\textit{Bias}, \textit{Cov}),$$

where "Dist" is a given distribution (e.g., Gaussian) with mean vector *Bias*, and covariance matrix *Cov*.

2. Figures of Merit (FOMs)

Traditional FOMs are bias (*Bias*), and mean squared prediction error (*Mspe*). In what follows, we propose functions of them that allow straightforward simultaneous inference on the whole state vector x.

Icv: Inverse of the coefficient of variation.

$$\textit{Icv} = \textit{Sdv}^{-1}\textit{Bias} = \textit{Sdv}^{-1}\{\mathrm{E}_y(\hat{x}(y) - \mathrm{E}_{x|y}(x))\},$$

Sdv: Square root of the diagonal matrix obtained from the prediction-error covariance matrix, *Cov*.

$$\textit{Sdv} = \{\mathrm{diag}(\textit{Cov})\}^{1/2} = \{\mathrm{diag}(\mathrm{E}_y(\mathrm{cov}_{x|y}(x)) + \mathrm{cov}_y(\hat{x}(y) - \mathrm{E}_{x|y}(x)))\}^{1/2},$$

Cor: Prediction-error correlation matrix.

$$\textit{Cor} = \textit{Sdv}^{-1}(\textit{Cov})\textit{Sdv}^{-1},$$

where we can show *Cov* to be equal to *Mspe*$-$(*Bias*)(*Bias*)$'$. From the simulation experiment, we obtain asymptotically unbiased estimates $\hat{Icv}$, $\hat{Sdv}$, and $\hat{Cor}$ of the FOMs *Icv*, *Sdv*, and *Cor* using,

$$\hat{Bias} \equiv \frac{1}{L}\sum_{l=1}^{L}(\hat{x}(y^{(l)}) - x^{(l)})$$

$$\hat{Mspe} \equiv \frac{1}{L}\sum_{l=1}^{L}(\hat{x}(y^{(l)}) - x^{(l)})(\hat{x}(y^{(l)}) - x^{(l)})'$$

$$\hat{Cov} \equiv \hat{Mspe} - (\hat{Bias})(\hat{Bias})'.$$

Then

$$\hat{Icv} = \hat{Sdv}^{-1}\hat{Bias}$$

$$\hat{Sdv} = \{\mathrm{diag}(\hat{Cov})\}^{1/2}$$

$$\hat{Cor} = \hat{Sdv}^{-1}(\hat{Cov})\hat{Sdv}^{-1}.$$

3. Simulations from a Bivariate State-Space Model

We use 20 simulated radiance measurements, y, to predict a bivariate state vector comprising the volume mixing ratios of CO_2 and O_2 in parts per million (ppm), $x = (x_1, x_2)'$, which is assumed to be generated from a bivariate Gaussian distribution. We predict the state vector, x, using Bayesian posterior analysis; specifically, $\hat{x}(y)$=posterior mode. For the simulation, we have:

- Prior distribution for the hidden state $x = (x_1, x_2)'$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim \text{Gau}\left(\mu_x = \begin{pmatrix} 390 \\ 209550 \end{pmatrix}, \Sigma_x = \begin{bmatrix} 4 & 8\rho \\ 8\rho & 16 \end{bmatrix}\right),$$

where ρ is the correlation between x_1 and x_2.

- Measurement error, $\epsilon_j^{(l)}$, where realisations are independently distributed as $\epsilon_j^{(l)} \sim \text{Gau}(0, \sigma_\epsilon^2)$, for j = 1,...,20.

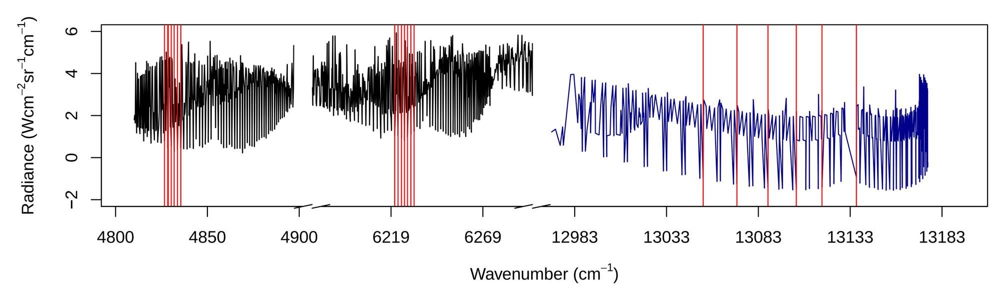
Radiances are measured for a notional vertical column of the atmosphere (4.5 – 6.5 km) in which pressure decreases approximately linearly with height. We assume aerosol-free nadir measurements at a representative height of 5.5 km, with constant temperature T_m = 252.43 K, pressure P_m = 0.4988 atm, solar flux, and reflectance; and we ignore changes in radiance due to emission and scattering effects. Now optimal estimation (i.e., the posterior mode) is used to solve the inverse problem and obtain a predictor $\hat{x}(y)$ for x.

3.1 The State-Space Model

To calculate the radiance at wavenumber ν_i, we used the parameters from the HITRAN2012 database (see Rothman et al., 2013) and calculated the line strength at temperature T_m for each wavenumber ν_i. Then, ignoring the self-broadened half-width, we used the air-broadened half-width at (T_m, P_m), the optical mass of state variable x_k, and the Lorentz-line-shape function (since altitude <16 km), to obtain $K_k(\nu_i, T_m, P_m)$, the absorption cross-section of x_k. Finally, using the Beer-Lambert law, we calculated the total radiance. Hence, for a spectral interval centred at ν_j that includes n_j wavenumbers $\{\nu_{jr} : r = 1, ..., n_j\}$, the nonlinear forward function $F_j(x)$ is given by,

$$F_j(x) = C_{IR}\exp\left(-\sum_{k=1}^{2} x_k \sum_{r=1}^{n_j} g(\nu_{jr}, T_m, P_m)\right),$$

where the constant C_{IR} approximates the solar flux and reflectance parameters, and $g(\nu_{jr}, T_m, P_m)$ represents the radiative transfer equation at wavenumber ν_{jr}. Finally, we selected J = 20 intervals to use in the experiment: 7 each from the CO_2 strong and weak bands, and 6 from the O_2 A-band, marked in red below.



We simulated L = 20,000 realisations of the state variables, $x^{(l)} = (x_1^{(l)}, x_2^{(l)})'$, l = 1,...,20,000, for different values of ρ and then defined different values of σ_ϵ^2 to obtain different signal-to-noise ratios (SNRs). Then, for each l we simulated a data vector of J = 20 radiances, $y^{(l)} = (y_1^{(l)}, ..., y_{20}^{(l)})'$, using

$$y_j^{(l)} = F_j(x^{(l)}) + \epsilon_j^{(l)}; \quad j = 1, ..., 20. \tag{1}$$

Finally, the posterior mode, $\hat{x}(y^{(l)})$ was obtained (Rodgers, 2000); l = 1,...,20,000.

3.2 Simulated FOMs

For the case, SNR = 0.5 and $\rho = -0.8$, we obtained from the simulation described above, the following:

$\hat{Icv}$	CO_2	O_2
	0.0022	-0.0143

$\hat{Bias}$	CO_2	O_2
	0.0018	-0.0390

$\hat{Cov}$	CO_2	O_2
CO_2	0.667	-1.068
O_2	-1.068	7.468

$\hat{Mspe}$	CO_2	O_2
CO_2	0.667	-1.068
O_2	-1.068	7.470

$\hat{Sdv}$	CO_2	O_2
CO_2	0.817	0
O_2	0	2.747

$\hat{Cor}$	CO_2	O_2
CO_2	1	-0.478
O_2	-0.478	1

The FOMs $\hat{Icv}$, $\hat{Sdv}$, and $\hat{Cor}$ can be used to compare properties of the prediction error for different SNR and ρ values.

4. Simultaneous Inference

Assume that "Dist" is *approximately* Gaussian, and define $c \equiv \chi_K^2(0.95)$ as the upper 95^{th} percentile of a chi-squared distribution on K (= 2) degrees of freedom. Then, using the Wald statistic, an approximate 95% *prediction ellipsoid* for the K-dimensional state x is defined by,

$$\textit{Ell}(0.95) \equiv \{x : (x - \hat{x}(y) + \textit{Bias})'\textit{Cov}^{-1}(x - \hat{x}(y) + \textit{Bias}) \leq c\}.$$

Now define $\omega \equiv (\textit{Cor})^{-1/2}\textit{Sdv}^{-1}x$ and correspondingly $\hat{\omega}(y) \equiv (\textit{Cor})^{-1/2}\textit{Sdv}^{-1}\hat{x}(y)$, and transform the ellipse to obtain the K-dimensional 95% *prediction spheroid*,

$$\textit{Sph}(0.95) \equiv \{\omega : (\omega - \hat{\omega}(y) + \textit{Cor}^{-1/2}\textit{Icv})'(\omega - \hat{\omega}(y) + \textit{Cor}^{-1/2}\textit{Icv}) \leq c\},$$

centred at $(\hat{\omega}(y) - \textit{Cor}^{-1/2}\textit{Icv})$ with radius $c = \chi_K^2(0.95)$ such that, approximately, $\Pr(\textit{Sph}(0.95)) = 0.95$.

For a single new realisation of the hidden 'true' state, x, where SNR = 0.5 and $\rho = -0.8$, we used the forward model (1) to generate a vector of radiances y, and we obtained a prediction $\hat{x}(y)$. The 95% prediction sphere, *Sph*(0.95), is easily calculated, and hence the 95% prediction ellipsoid, *Ell*(0.95), is obtained by back-transformation.

Alternatively, a 95% univariate prediction interval for each state variable (e.g., x_1) is given by,

$$\Pr\left(\frac{|\hat{x}_1 - x_1 - \textit{Bias}_1|}{(\textit{Cov}_{11})^{1/2}} < c_u\right) = 0.95,$$

where c_u is the critical value of a univariate Gaussian distribution. A univariate prediction interval ignores the presence of other state variables, and here

$$\Pr\left(\left\{\frac{|\hat{x}_1 - x_1 - \textit{Bias}_1|}{(\textit{Cov}_{11})^{1/2}} < c_u\right\} \cap \left\{\frac{|\hat{x}_2 - x_2 - \textit{Bias}_2|}{(\textit{Cov}_{22})^{1/2}} < c_u\right\}\right) = 0.909,$$

which is less than 0.95.

A nominal Bonferroni-adjusted 95% simultaneous prediction region for x satisfies,

$$\Pr\left(\left\{\frac{|\hat{x}_1 - x_1 - \textit{Bias}_1|}{(\textit{Cov}_{11})^{1/2}} < c_b\right\} \cap \left\{\frac{|\hat{x}_2 - x_2 - \textit{Bias}_2|}{(\textit{Cov}_{22})^{1/2}} < c_b\right\}\right) \geq 0.95,$$

where c_b is the Bonferroni-adjusted critical value for a univariate Gaussian distribution; here $c_b = (1.1436)c_u$. The prediction region is larger than the region formed by the univariate prediction intervals. Here, its probability is 0.953, which is greater than 0.95, illustrating that the Bonferroni-adjusted prediction region is conservative.

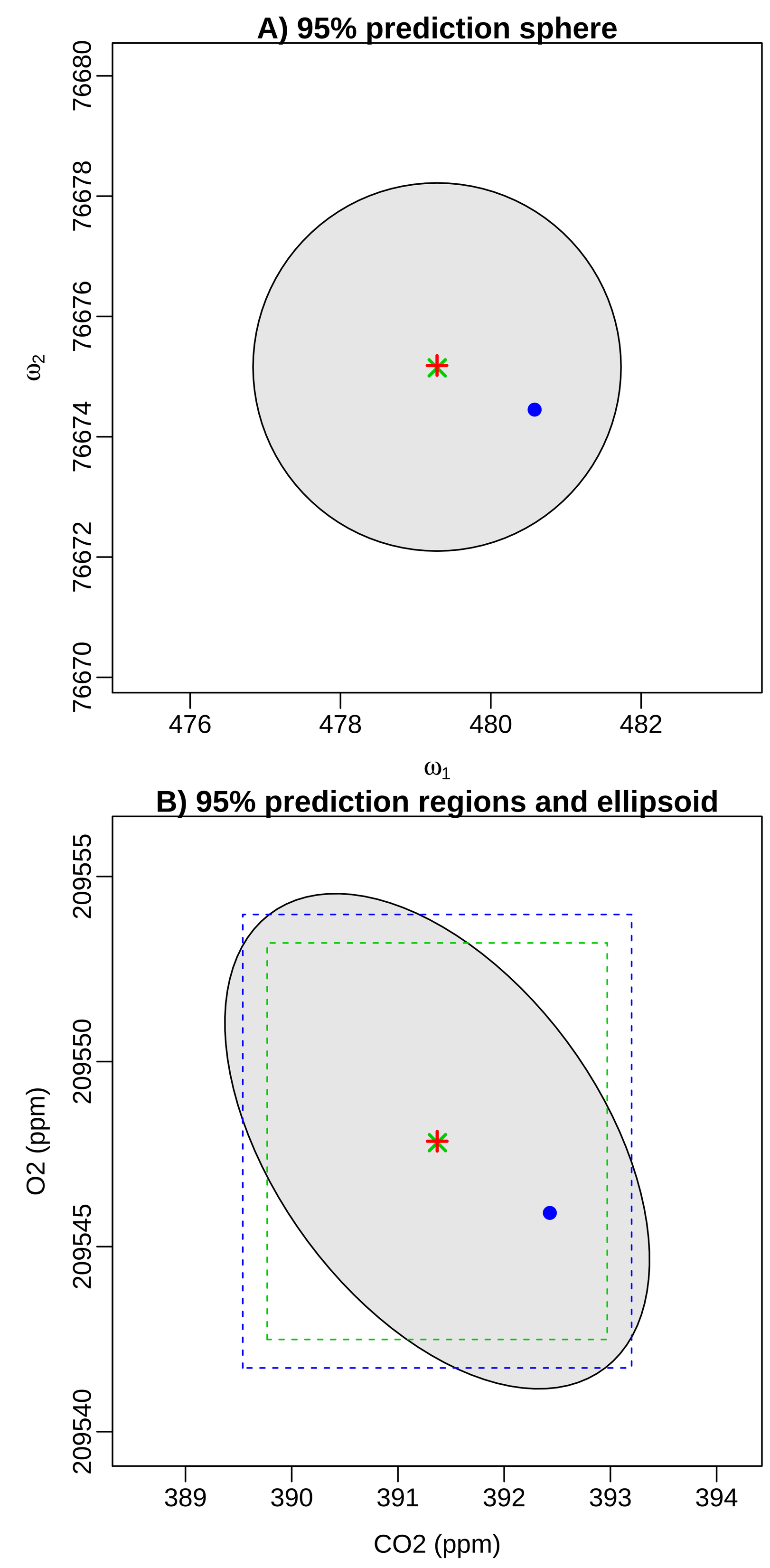


Figure 2: *SNR = 0.5 and $\rho = -0.8$. Key: × prediction; + bias; ······ Region from univariate prediction intervals; ······ Bonferroni-adjusted prediction region; • true state.*

Of course, in practice, we do not know the true value (•).

Conclusions: These results illustrate that simultaneous inference on the state x is more efficient than inferring individual state elements one-at-a-time, and we see that FOMs *Icv*, *Sdv*, and *Cor* have an easily interpretable role in simultaneous inference.

References

Rodgers, CD (2000), *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, Singapore.

Rothman, LS, Gordon, IE, Babikov, Y et al. (2013), 'The HITRAN2012 molecular spectroscopic database,' *Journal of Quantitative Spectroscopy and Radiative Transfer*, **130**, pp. 4–50.